Predictions of Crystal Structures from First Principles

Rafał Podeszwa and Krzysztof Szalewicz
Department of Physics and Astronomy, University of Delaware, Newark, DE
poszwa@physics.udel.edu and szalewic@udel.edu

Robert Bukowski
Cornell Theory Center, Cornell University, Ithaca, NY
rb299@cornell.edu

Betsy M. Rice
US Army Research Laboratory (ARL), Aberdeen Proving Ground, MD
betsyr@arl.army.mil

Abstract

A recently developed method denoted as SAPT(DFT), which applies symmetry-adapted perturbation theory (SAPT) based on Kohn-Sham orbitals and orbital energies and includes the dispersion component obtained using frequency-dependent density susceptibilities from density functional theory (DFT), has been shown to provide as accurate interaction energies as high-level wave-function-based methods. At the same time, the former calculations can be performed at a greatly reduced computational cost compared to the latter, in fact, in a time comparable to supermolecular DFT calculations. The SAPT(DFT) method is particularly important for systems with a dominant dispersion component since the supermolecular DFT approach fails completely in this case.

SAPT(DFT) was used to compute the interaction potential for the RDX dimer. This potential was applied to predictions of the properties of the RDX crystal in molecular dynamics simulations. The fully ab initio calculated properties are in excellent agreement with experiment and the predictions are even slightly better than achieved by empirical potentials fitted to the crystal experimental data.

1. Motivation

In 1988, Maddox published in Nature[1] a provocative op-ed stating that "one of the continuing scandals" is that computational scientists are not able to predict crystal structures from molecular structures. This opinion was echoed in the same journal first in 1996 by Ball[2] and more recently by Desiraju[3], who in 2002 wrote that the issue "eluded scientists for more than 50 years". Desiraju pointed out in particular to low success rate of predictions in the blind tests conducted by the Cambridge Crystalographic Data Center (CCDC)[4,5]. The situation does not seem to improve as the success rate of the third test[5] was lower than that of the previous ones. This is indeed an unsatisfactory situation in view of the importance of such crystals. For example, energetic materials are crystals of large organic molecules and polymorphism of drugs is a major problem in pharmaceutical industry.

Maddox's opinion was partly due to a misconception that crystal structures are simple functions of molecular structures. We now know that this is not the case and the crystal structures depend in a subtle way on intermolecular force fields. Such force fields can be computed *ab initio* using wave-function (WF) based methods, but until recently the accuracy of such calculations for molecules containing more than a few atoms was far from quantitative and was insufficient for determination of crystal structures. One might have hoped that the problem could be resolved by the development of density functional theory (DFT) which can be applied to systems containing hundreds of atoms, but this method turned out to fail badly for all intermolecular interactions except those involving highly polar species. Recently, a method has been proposed[6–12] which combines symmetry-adapted perturbation theory (SAPT)[13,14] of intermolecular interactions with the DFT representation of monomers. The efficiency of SAPT(DFT) has been improved[7,15–17] by applying the density-fitting method[18]. For the current capabilities of SAPT(DFT), and references to earlier work, see Reference 15. The predictions of SAPT(DFT) are as accurate as those of high-level WF-based electronic structure methods, whereas, at the same time, for systems within the current range of applicability, a SAPT(DFT) calculation takes in fact less time than a supermolecular DFT calculation. SAPT(DFT) has been applied to

0-7695-3088-5/07 $25.00 © 2007 IEEE

compute the complete potential surfaces of the water[19] and benzene dimers[20], in each case giving results in excellent agreement with experimental data. Therefore, one may hope that the force fields computed by SAPT(DFT) will be accurate enough to predict crystal structures.

The present paper tests the ability of the force fields computed by SAPT(DFT) to predict the crystal structure of cyclotrimethylene trinitramine ($C_3N_6O_6H_6$), known also under name of RDX. This molecule is one of the most important energetic materials. It has been the subject of many theoretical investigations[21-35]. The dimer of RDX is bound by van der Waals forces with a very significant contribution from the dispersion forces[36]. The dispersion forces are particularly challenging for WF-based methods since expensive, high-level approaches—such as for example the coupled-cluster method with single, double, and non-iterative triple excitations [CCSD(T)]—are required to predict reasonably accurate interaction energies. The large dispersion contribution is also the reason why DFT methods are unable to predict properties of such systems. For example, when DFT methods were applied to the RDX crystal in Reference 32, the predicted density differed by about 12% from experiment. Since the RDX dimer is too large for advanced WF-based methods, even for single-point calculations, this system was computationally intractable until the development of SAPT(DFT). This dimer, apart from the interest in the system itself, is a good proving ground for the performance of the SAPT(DFT) method on a significantly larger systems than the benzene dimer, the largest dimer to date for which an ab initio potential energy surface (PES) has been developed[20]. By the number of electrons, RDX is almost three times larger than benzene. Thus, a CCSD(T) calculation for the RDX dimer would be three orders of magnitude more time consuming than for the benzene dimer. The effect of this is that whereas it was possible to perform single-point CCSD(T) calculations for the benzene dimer with basis sets including diffuse functions for selected configurations, the RDX dimer is completely beyond the reach of CCSD(T), even with the largest possible supercomputer resources. This also means that the only validations of theoretical work on RDX come from comparisons to experimental results. The work reviewed in this paper is described in more detail in Reference 37.

2. Computational Details

For our calculations, we have used the density-fitting implementation of SAPT(DFT) described in detail in Reference 15. The DFT calculations for the monomers were performed using the dalton[38] program. The Kohn-Sham orbitals of the monomer were obtained using the PBE0 functional[39,40] with the asymptotic correction of Grüning, et al.[41]. The ionization potential needed for the asymptotic correction, equal to 0.373 hartree, was taken from Reference 15. The monomer geometry has been taken from the experimental crystal structure[42], but "symmetrized" by averaging the positions of the atoms with the mirror images of the corresponding atoms obtained by reflection in the pseudo-mirror plane. We have used the aug-cc-pVDZ basis set of Kendall, et al.[43] in the MC$^+$BS approach[44]. This basis set was further supplemented with midbond functions from Reference 15. The auxiliary basis set needed for density fitting was taken from References 45 and 46.

We have first produced a tentative fit to a small number of *ab initio* points obtained for radial cross-sections with angular orientations from several nearest-neighbor dimer structures in the crystal. A regular grid of *ab initio* points was then developed, with angular parameters spaced every 90° and radial distances chosen to evenly cover the tail, well, and potential wall in radial cross-sections for each orientation of monomers.

3. Analytic Fit

The potential surface was fitted to a site-site formula

$$V = \sum_{a\in A}\sum_{b\in B} u_{ab}\left(r_{ab}\right), \qquad (1)$$

where the summation runs over all atomic sites (positions of nuclei) a of monomer A, sites b of monomer B, and r_{ab} denotes the distance between two such sites. The function u_{ab} given by

$$u_b = \left(1 + a_1^{ab} r_{ab} + a_2^{ab} r_{ab}^2\right) e^{\alpha_a - \beta_a r_{ab}} + f_1\left(\delta_1^{ab}, r_{ab}\right)\frac{q_a q_b}{r_{ab}} + f_6\left(\delta_6^{ab}, r_{ab}\right)\frac{C_6^{ab}}{r_{ab}^6} \qquad (2)$$

is a simplification of the function used for the benzene dimer[20]. To alleviate the $1/r_{ab}^n$ divergent character of the two latter terms in Eq. (2) at short intermolecular distances, these terms include the Tang-Toennies damping functions[47]

$$f_n\left(\delta, r\right) = 1 - e^{-\delta r}\sum_{m=0}^{n}\frac{\left(\delta r\right)^m}{m!}. \qquad (3)$$

The charges q_a and q_b modeling the electrostatic interactions and the C_{ab}^6 van der Waals coefficients modeling the long-range dispersion and induction interactions were obtained from separate fits of these interactions to independent asymptotic expansions at very large intermonomer separations. The latter expansions were obtained from center-of-mass (COM) monomer multipole moments and polarizabilities computed *ab initio* at the same level of theory as our calculations at

finite separations. In this way, the fit nearly exactly reproduces SAPT(DFT) points at large separations. The remaining parameters were fitted by a least-square method. The root-mean-square deviation (RMSD) of the fit calculated for points with $E_{int} < 0$ was 0.13 kcal/mol, whereas the overall RMSD was 0.47 kcal/mol.

4. Minima on the Potential Energy Surface

The fitted PES of the dimer was explored with a simple implementation of the eigenvector-following local minimization method[48] using 18,000 starting points with randomly selected angular configurations and randomly selected COM separation R between 4 and 10 Å. We have found 54 minima on the PES.

The energies of the minima range between −12.5 and −5.0 kcal/mol and the distances between COM of monomers between 4.25 and 7.24 Å. The structure of the global minimum is presented as the first dimer in Figure 1. This structure has a fairly large R equal to 6.69 Å, near the largest ones among all the minima. It is bound by strong electrostatic forces due to interactions of the equatorial NO_2 groups with the axial hydrogens on the interacting partner, a hydrogen-bond interaction with the shortest O-H distance of 2.58 Å. The second-lowest structure of energy equal to −10.16 kcal/mol (the second dimer in Figure 1) also includes hydrogen bonds, with the shortest O-H distance equal to2.33 Å and2.75 Å and two longer ones equal to 2.61 Å.

Among many other classes of minima, an important one is characterized by very short inter-monomer distances. The structure shown as the third dimer in Figure 1, with R = 4.50 Å and energy equal to −7.95 kcal/mol, is the lowest-energy member of this class. This structure correspond to a slipped parallel orientation of the $(CN)_3$ rings with the antiparallel orientation of the axial NO_2 groups. Although the electrostatic energy is negative, the "dispersionless" energy is strongly positive and the structure achieves the majority of stability due to the dispersion interaction. This minimum is very close geometrically and energetically to the nearest-neighbor dimer structure in the RDX crystal.

5. Physical Decomposition of the Force Field

Since SAPT computes the physical contributions to the interaction energy, it is worthwhile to analyze the relative importance of these contributions, as it was done already in some detail in Section 4. More extensive analysis is presented in Figures 2 and 3. These figures also show the accuracy of the fit and a comparison to the SRT empirical potential[23].

The two lowest-energy structures (the consecutive graphs in Figure 2) represent similar physical pictures. The attractive contribution is dominated by the electrostatic energy in the whole range of R. The relatively slow decay of the electrostatic energy with R is a result of fairly strong dipole-dipole interactions. Nevertheless, the dispersion energy is also quite important and in the regions close to the minimum it is almost equal to the electrostatic energy. For these structures, the induction energy components are much less important, in particular for the second lowest minimum. Moreover, a large part of the induction energy (except at long range) is quenched by the exchange-induction energy: for the minimum, more than half of the induction energy is quenched by the latter term. For both cross-sections, the empirical SRT potential[23] is very close to SAPT(DFT) at long-range (this is the result of accurate charges obtained from second-order Møller-Plesset (MP2) level calculations used in the SRT potential), while near the minimum and at the short range the SRT curves are above SAPT(DFT) [about 2 kcal/mol at SAPT(DFT) minimum] and the minimum position is shifted to a longer R by about 0.15 Å.

The third structure, shown in Figure 3, has a significantly different character than the two structures describes above. It corresponds to a minimum with the second shortest intermolecular distance, more than 2 Å shorter than that of the global minimum. The attractive contribution is clearly dominated by the dispersion component for the whole range of intermolecular separations. The electrostatic interaction is quite small. The considered angular orientation is in fact dipole-dipole repulsive since, as seen in Figure 3, at larger intermolecular separations, where the electrostatic energy is dominated by the slowest-decaying dipole-dipole contribution, the electrostatic energy is positive. The induction energy is of similar magnitude to the electrostatic energy (in contrast to the first two minima), but it is even more strongly quenched (about 80% at the minimum distance) than in the previously described configurations. This angular configuration, despite a very different character of the monomers, is surprisingly similar to the slipped-parallel configuration of the benzene dimer (although here the repulsive electrostatic energy is caused by the dipole-dipole interaction instead of the quadrupole-quadrupole interaction), where a similar mutual interplay of the various terms has been observed[20]. The latter structure was also the minimum for the benzene dimer with the shortest intermolecular distance among all the minima. For the structure of Figure 3, the SRT potential[23] is astonishingly close to the SAPT(DFT) potential, much closer than for the two orientations from Figure 2. This is likely due to the fact that the SRT potential was fitted to the properties of the RDX crystal and the structure of Figure 3 is the major dimer configuration in the crystal. To provide a broader comparison between SAPT(DFT) and the SRT potential,

we calculated the RMSD of the SRT potential on the set of *ab initio* SAPT(DFT) points. The deviations amounted to 2.85 and 1.28 kcal/mol for all points and the points with negative energies, respectively.

6. Molecular Dynamics Simulations of RDX Crystal

The PES fit described in previous sections has been applied in molecular dynamics simulations of the RDX crystal. We performed isothermal-isobaric (NPT), single-trajectory simulations at 298K with 1 bar pressure. The trajectory was integrated for 130 ps with 0.001 ps timestep. During the first 30,000 steps (30ps), the structure was equilibrated. The simulation cell size was 3×3×3 unit cells and the starting configurations were taken from the experimental geometry. We have also used here the exact experimental monomer geometry instead of the symmetrized one used in *ab initio* calculations and in the fitting process, but the parameters of the fit were kept unchanged. During the simulation, the monomers were assumed rigid. The calculations were performed with the DLPOLY program[49]. The cut-off used was equal to 12 Å. Long-range corrections for the energy and virial due to the dispersion and induction contributions were calculated using a method described in Reference 50. The long-range Coulomb interaction energy was corrected by Ewald summation technique[50].

Table 1. Properties of the RDX crystal: energies per molecule (kJ/mol), density (g/cm^3), and cell vectors (Å)

	Energy	Density	Cell vectors		
SAPT(DFT)	−120.21	1.784	13.287	11.633	10.701
empirical (SRT)[a]	−115.22	1.738	13.404	11.799	10.732
experiment	−130.1[b]	1.806[c]	13.182[c]	11.574[c]	10.709[c]

[a] T = 298 K.
[b] Experimental sublimation energy from Ref. 51
[c] Ref. 42

The results of the simulations are presented in Table 1, together with those from simulations with the empirical potential of Reference 23. The simulated crystal structure is in an excellent agreement with experiment. The density is only 1.2% too low, and the *a*, *b*, and *c* vector units have errors of 0.8%, 0.5%, and −0.1%, respectively. It is encouraging that the *ab initio* results are even more accurate than those given by the empirical SRT potential of Reference 23, which was fitted to the experimental RDX crystal data.

7. Summary and Conclusions

We have developed a complete 6-dimensional potential energy surface for the RDX dimer. All this development was based completely on first principles, i.e., without any use of experimental results (except that the geometry of monomers was derived from X-ray measurement, but an ab initio minimization of the geometry would give a very similar structure). This system, due to its size, is intractable by any high-level ab initio method, but SAPT(DFT). The complicated character of the potential energy surface required a fit with a large number of parameters, which in turn required a large number, more than one thousand, of single-point calculations.

We have found 54 minima on the PES of the system. The large number of minima is a result of the complicated structure of the monomer. The minima were analyzed in terms of physical components of the interaction energy. We found that for the RDX dimer the dispersion energy is overall the most important component. Although the electrostatic interaction is often also an important part of the interaction energy (in a few cases it is actually the most important attractive contribution), the dispersion contribution is always very important. Since the dispersion interaction cannot be properly modeled by the standard supermolecular DFT, the current results show that the poor performance of the supermolecular DFT in predicting the RDX crystal structure in Reference 32 was indeed caused by the large dispersion contributions in this system.

Molecular dynamics simulations of the RDX crystal structure performed using the SAPT(DFT) potential yielded an excellent agreement with the experiment. The density of the crystal obtained from the simulations is only about 1.2% lower than the experimental density. This result is significantly better than the density obtained from simulations with the SRT empirical potential[23].

Acknowledgments

Funding for this work was provided by an Army Research Office DEPSCOR grant. Most of calculations were performed using computer resources granted by a DoD HPCMP Challenge Project.

References

1. Maddox, J., *Nature*, 335, 201, 1988.

2. Ball, P., *Nature*, 381, 648, 1996.

3. Desiraju, G.R., *Nature Materials*, 1, 77, 2002.

4. Motherwell, W.D.S., H.L. Ammon, J.D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D.W.M. Hofmann, F.J.J.

Leusen, J.P.M. Lommerse, W.T.M. Mooij, et al., *Acta Crystallogr. B*, 58, 647, 2002.

5. Day, G.M., W.D.S. Motherwell, H.L. Ammon, S.X.M. Boerrigter, R.G. Della Valle, E. Venuti, J.D. Dunitz, B. Schweizer, B.P. van Eijck, P. Erk, et al., *Acta Crystallogr. B*, 61, 511, 2005.

6. Misquitta, A.J. and K. Szalewicz, *Chem. Phys. Lett.*, 357, 301, 2002.

7. Misquitta, A.J., B. Jeziorski, and K. Szalewicz, *Phys. Rev. Lett.*, 91, 033201, 2003.

8. Misquitta, A.J. and K. Szalewicz, *J. Chem. Phys.*, 122, 214109, 2005.

9. Misquitta, A.J., R. Podeszwa, B. Jeziorski, and K. Szalewicz, *J. Chem. Phys.*, 123, 214103, 2005.

10. Hesselmann, A. and G. Jansen, *Chem. Phys. Lett.*, 357, 464, 2002.

11. Hesselmann, A. and G. Jansen, *Chem. Phys. Lett.*, 367, 778, 2003.

12. Williams, H.L. and C.F. Chabalowski, *J. Phys. Chem. A*, 105, 646, 2001.

13. B. Jeziorski, R. Moszyński, and K. Szalewicz, *Chem. Rev.*, 94, 1887, 1994.

14. Szalewicz, K., K. Patkowski, and B. Jeziorski, in *Intermolecular Forces and Clusters*, edited by D.J. Wales, Springer, 2005, Structure and Bonding, pp. 1–72.

15. Podeszwa, R., R. Bukowski, and K. Szalewicz, *J. Chem. Theory Comput.*, 2, 400, 2006.

16. Hesselmann, A., G. Jansen, and M. Schütz, *J. Chem. Phys.*, 122, 014103, 2005.

17. Bukowski, R., R. Podeszwa, and K. Szalewicz, *Chem. Phys. Lett.*, 414, 111, 2005.

18. Dunlap, B.I.,J.W.D. Connolly, and J.R. Sabin, *J. Chem. Phys.*, 71, 4993, 1979.

19. Bukowski, R., K. Szalewicz, G. Groenenboom, and A. van der Avoird, *J. Chem. Phys.*, 125, 044301, 2006.

20. Podeszwa, R., R. Bukowski, and K. Szalewicz, *J. Phys. Chem. A*, 110, 10345, 2006.

21. Kunz, A.B., *Phys. Rev. B*, 53, 9733, 1996.

22. Harris, N.J. and K. Lammertsma, *J. Am. Chem. Soc.*, 119, 6583, 1997.

23. Sorescu, D.C., B.M. Rice, and D.L. Thompson, *J. Phys. Chem. B*, 101, 798, 1997.

24. Rice, B.M. and C.F. Chabalowski, *J. Phys. Chem. A*, 101, 8720, 1997.

25.. Sorescu, D.C., B.M. Rice, and D.L. Thompson, *J. Phys. Chem. B*, 103, 6783, 1999.

26. Kuklja, M.M. and A.B. Kunz, *J. Appl. Phys.*, 86, 4428, 1999.

27. Kuklja, M.M. and A.B. Kunz, *J. Phys. Chem. Solids*, 61, 35, 2000.

28. Chakraborty, D., R.P. Muller, S. Dasgupta, and W.A. Goddard, *J. Phys. Chem. A*, 104, 2261, 2000.

29. Kuklja, M.M., *J. Phys. Chem. B*, 105, 10159, 2001.

30. Vladimiroff, T. and B.M. Rice, *J. Phys. Chem.*, A 106, 10437, 2002.

31. Perger, W.F., R. Pandey, M.A. Blanco, and J. Zhao, *Chem. Phys. Lett.*, 388, 175, 2004.

32. Byrd, E.F.C., G.E. Scuseria, and C.F. Chabalowski, *J. Phys. Chem. B*, 108, 13100, 2004.

33. Agrawal, P.M., B.M. Rice, L.Q. Zheng, and D.L. Thompson, *J. Phys. Chem. B*, 110, 26185, 2006.

34.. Zheng, L.Q and D.L. Thompson, *J. Chem. Phys.*, 125, 084505, 2006.

35. Byrd, E.F.C. and B.M. Rice, *J. Phys. Chem. C*, 111, 2787, 2007.

36. Podeszwa, R.R., Bukowski, and K. Szalewicz, in *Proceedings of the HPCMP User Group Conference, IEEE, 2006*, pp. 189–192.

37. Podeszwa, R.R. Bukowski, B.M. Rice, and K. Szalewicz, *Phys. Chem. Chem. Phys.*, 2007, submitted.

38. *DALTON, a molecular electronic structure program, Release 2.0*, see http://www.kjemi.uio.no/software/dalton/dalton.html, 2005.

39. Perdew, J.P., K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.*, 77, 3865, 1996.

40. Adamo, C. and V. Barone, *J. Chem. Phys.*, 110, 6158, 1999.

41. Grüning, M., O.V. Gritsenko, S.J.A. van Gisbergen, and E.J. Baerends, *J. Chem. Phys.*, 114, 652, 2001.

42. Choi, C.S. and E. Prince, *Acta Crystallogr.*, B28, 2857, 1972.

43. Kendall, R.A., T.H. Dunning, and R.J. Harrison, *J. Chem. Phys.*, 96, 6796, 1992.

44. Williams, H.L., E.M. Mas, K. Szalewicz, and B. Jeziorski, *J. Chem. Phys.*, 103, 7374, 1995.

45. Weigend, F., A. Köhn, and C. Hättig, *J. Chem. Phys.*, 116, 3175, 2002.

46. Weigend, F., *Phys. Chem. Chem. Phys.*, 4, 4285, 2002.

47. Tang, K.T. and J.P. Toennies, *J. Chem. Phys.*, 80, 3726, 1984.

48. Cerjan, C.J. and W.H. Miller, *J. Chem. Phys.*, 75, 2800, 1981.

49. Smith, W. and T.R. Forester, *J. Mol. Graph.*, 14, 136, 1996.

50. Allen, M.P. and D.J. Tindesley, *Computer Simulation of Liquids*, Oxford University Press, New York, 1989.

51. Rosen, J.M. and C. Dickinson, *J. Chem. Eng. Data*, 14, 120, 1969.

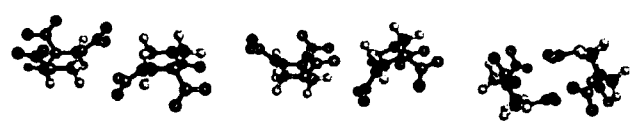

Figure 1. M1, M2, and M3 structures

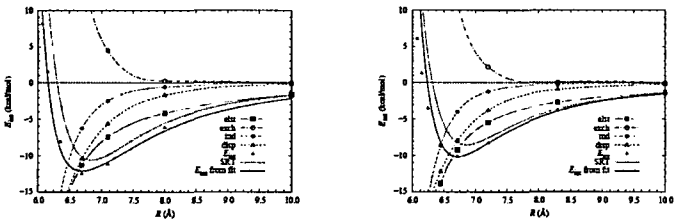

Figure 2. Radial cross section of the angular configuration of structures M1 and M2: 'elst'—electrostatic energy, 'exch'—sum of the first and second-order exchange energies, 'ind'—induction energy, 'disp'—dispersion energy, 'SRT'—potential from Reference 23

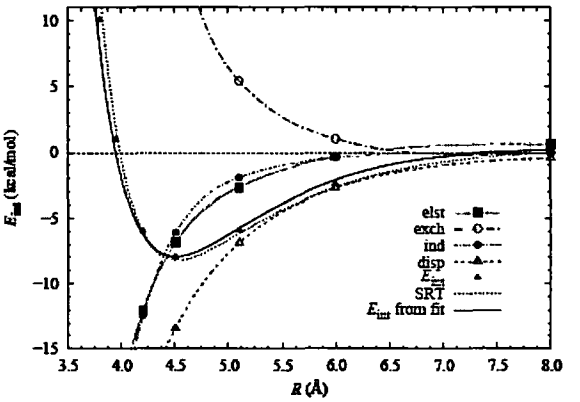

Figure 3. Radial cross section of the angular configuration of structure M3. For description of symbols, see Figure 2.